



LAWRENCE  
LIVERMORE  
NATIONAL  
LABORATORY

LLNL-TR-681077

# Arbitrary Dimension Convection-Diffusion Schemes for Space-Time Discretizations

R. E. Bank, P. S. Vassilevski, L. T. Zikatanov

January 20, 2016

## **Disclaimer**

---

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

# Arbitrary Dimension Convection-Diffusion Schemes for Space-Time Discretizations<sup>☆</sup>

Randolph E. Bank

*Department of Mathematics, University of California, San Diego, La Jolla, CA 92093.*

*Email:rbank@ucsd.edu*

Panayot S. Vassilevski\*

*Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, P.O.*

*Box 808, Mail Stop L-561, Livermore, CA 94551. Email:panayot@llnl.gov*

Ludmil T. Zikatanov

*Department of Mathematics, The Pennsylvania State University, University Park,  
PA 16802. Email:ludmil@psu.edu*

*Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 1113 Sofia,  
Bulgaria.*

---

## Abstract

This note proposes embedding a time dependent PDE into a convection-diffusion type PDE (in one space dimension higher) with singularity, for which two discretization schemes, the classical streamline-diffusion and the EAFE (edge average finite element) one, are investigated in terms of stability and error analysis. The EAFE scheme, in particular, is extended to be arbitrary order which is of interest on its own. Numerical results, in combined space-time domain demonstrate the feasibility of the proposed approach.

*Keywords:* space-time discretization and solvers

---

## 1. Introduction

The embedding of time-dependent problems into a one space dimension higher stationary problem is not a new idea. It has many appealing properties, such as: using already existing tools developed for stationary problems;

---

<sup>☆</sup>This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The work of the first author was supported in part by the National Science Foundation under contract DMS-1318480 and the work of the third author was supported in part by the National Science Foundation under contracts DMS-1418843 and DMS-1522615.

\*Corresponding Author

using adaptive methods with reliable and efficient error control; the ability to use existing efficient solver libraries developed for stationary problems. There is, however, a drawback: typically, the memory needed to run a simulation using the combined space-time discretization approach is increased by an order of magnitude. One way to keep the memory required by such methods under control is to use time intervals with fixed length. Another, more general, remedy to the extensive use of computer memory in space-time simulations is to employ accurate dimension reduction algorithms, both in space and in time, which can lead to coarser problems with fewer degrees of freedom, also known as upscaled discretizations. Indeed, an accurate coarser problem can replace the expensive, in terms of memory, fine-grid one and still provide a reliable discretization tool. For a general dimension reduction approach by coarsening (in three space dimensions), we refer to [1]. The extension of the technique proposed in [1] to 4D space-time elements is a work in progress. Another feasible approach for dimension reduction in space-time discretizations is to exploit sparse grids, as proposed in [2]. More recently, discrete space-time schemes using B-splines and Non-Uniform Rational Basis Splines (NURBS) have been employed (see [3]) to yield stable isogeometric analysis methods for the numerical solution of parabolic PDEs in fixed and moving spatial domains.

We point out that in the present note we do not consider dimension reduction techniques. Rather, as a first step, we study the accuracy and stability of the proposed embedding. More specifically, for the discretization of the space-time formulation of a parabolic problem we exploit two well-known techniques for convection diffusion equations: the streamline diffusion method [4] (see also [5], [6]) and the EAFE–Edge Average Finite Element scheme [7] (see also [8] and [9]). Let us add that the high order EAFE method developed here provides a novel, high order, exponentially fitted discretization for convection-diffusion problems with suitable stability and approximation properties.

The structure of the remainder of this note is as follows. In Section 2, we introduce the space-time formulation of parabolic problems. Then, in Section 3, we present the streamline diffusion method in our space-time setting. Section 4, contains the derivation of the high order EAFE scheme on simplicial finite element grids in arbitrary spatial dimension. The application details for the lowest order EAFE discretization to parabolic problems is given in Section 5. Finally, in Section 6, we present numerical tests showing the optimality and efficiency of both schemes for space-time formulation of parabolic problems. We conclude this paragraph with remark on the terminology: as the EAFE scheme may be viewed as a multidimensional Scharfetter-Gummel discretization [10], in what follows, we use the terms “EAFE discretization” and “Scharfetter-Gummel discretization” interchangeably.

## 2. Space time formulation of parabolic problems

We consider the following parabolic problem:

$$\begin{aligned} u_t - \operatorname{div}(K(x)\nabla u - \mathbf{b} \cdot u) + \gamma u &= f, & x \in \Omega_s, \\ u &= 0, & x \in \Gamma = \partial\Omega_s; \quad u(x, 0) = u_0(x), & x \in \Omega_s. \end{aligned} \tag{1}$$

Here,  $\beta$  is a vector field (a velocity) and  $K(x)$  is, in general, a scalar (or  $d \times d$  tensor valued) function. Let  $\Omega_t = (0, t_{\max})$  be the time interval of interest. The space-time domain is  $\Omega = \Omega_s \times \Omega_t$ . For convenience we have assumed homogeneous Dirichlet boundary conditions  $u = 0$  on  $\partial\Omega_s \times \Omega_t$ . In treating time as a space-like variable, the initial condition at  $t = 0$  becomes a Dirichlet boundary condition for the  $(d+1)$  dimensional problem.

In a space-time formulation, introducing a new variable  $y = (x, t)$  then gives the following convection diffusion equation: Find  $u = u(y)$  such that

$$\begin{aligned} -\operatorname{div}_y(D\nabla_y u + \beta \cdot \nabla_y u) + \gamma u &= f \text{ in } \Omega \quad \mathbf{b} = (\beta^t, 1)^t : \Omega \mapsto \mathbb{R}^{d+1}, \\ u &= 0 \text{ on } \Gamma = \partial\Omega \times \Omega_t; \quad u = u_0 \text{ on } \Gamma_0 = \Omega_s \times \{t = 0\}. \end{aligned}$$

Without loss of generality we may assume that  $u_0 = 0$  and we define  $\mathcal{H}_E^1(\Omega)$  as the subspace of  $\mathcal{H}^1(\Omega)$  satisfying these homogeneous Dirichlet boundary conditions.

In the following we consider two schemes for discretization of convection diffusion problems and apply them to space-time formulations of (1). These are the Streamline Diffusion and the Scharfetter-Gummel (EAFE) discretizations.

For the latter we need a non-singular  $D$ , while above  $D = \begin{bmatrix} K & 0 \\ 0 & 0 \end{bmatrix}$  is actually degenerate. TO remedy this, we perturb it to make it invertible, i.e., we let  $D = \begin{bmatrix} K & 0 \\ 0 & \epsilon \end{bmatrix}$  for a small parameter  $\epsilon > 0$ .

### 3. Streamline Diffusion

We first consider a simple case when  $K = \alpha I$ ,  $\alpha > 0$ ,  $\gamma \geq 0$ , and  $\beta$  are constant. Then equation (1) has the form:

$$Lu \equiv u_t - \alpha \Delta u + \beta \cdot \nabla u + \gamma u = f \quad (2)$$

The results below generalize to the variable coefficient case in a straightforward and well-studied fashion. Here we consider the constant coefficients case only in an attempt to keep the focus on the important aspect of time discretization. In allowing for different sizes of  $\alpha$ ,  $\beta$  and  $\gamma$ , our analysis covers several scenarios of interest. For simplicity we assume the initial condition  $u_0 = 0$ .

The weak form of (2) is given by: find  $u \in \mathcal{H}_E^1$  such that

$$(u_t, v) + \alpha(\nabla u, \nabla v) + (\beta \cdot \nabla u, v) + \gamma(u, v) = f(v)$$

for all  $v \in \mathcal{H}_E^1$ . The space-time bilinear for  $B(u, v)$  is given by

$$B(u, v) = \int_0^T (u_t, v) + \alpha(\nabla u, \nabla v) + (\beta \cdot \nabla u, v) + \gamma(u, v) dt$$

where

$$(u, v) = \int_{\Omega_s} uv dx$$

is the usual  $\mathcal{L}_2$  inner product on  $\Omega_s$ . The right hand side is given by the linear functional

$$F(v) = \int_0^T f(v) dt.$$

The energy norm for this problem is given by

$$\|u\|^2 = \|u(T)\|^2 + \int_0^T \alpha \|\nabla u\|^2 + h\nu \|\beta \cdot \nabla u + u_t\|^2 + \gamma \|u\|^2 dt$$

where

$$\nu = \frac{1}{\sqrt{|\beta|^2 + 1}}.$$

We make a standard Petrov-Galerkin streamline diffusion discretization for this  $d+1$  dimensional problem. We assume that the space-time domain  $\Omega$  is covered by a shape regular quasiuniform tessellation  $\mathcal{T}_h$  of elements of size  $h$ . Let  $V_h \subset \mathcal{H}_E^1$  denote a  $C^0$  conforming piecewise polynomial finite element space. The space  $V_h$  itself is the trial space. In our Petrov-Galerkin formulation, the test functions are given by  $v + \theta h\nu(\beta \cdot \nabla v + v_t)$  for  $v \in V_h$ , where  $\theta$  is a parameter that will be characterized below. The discrete problem is: find  $u_h \in V_h$  such that

$$B_h(u_h, v) \equiv B(u_h, v) + \int_0^T (Lu_h, \theta h\nu(\beta \cdot \nabla v + v_t)) dt = F(v + \theta h\nu(\beta \cdot \nabla v + v_t))$$

for all  $v \in V_h$ . Because  $V_h$  is only  $C^0$ , the term  $(Lu_h, \theta h\nu(\beta \cdot \nabla v + v_t))$  is formally interpreted elementwise due to possible discontinuities on inter-element boundaries.

We begin with a basic stability result.

**Lemma 1.** *For  $v \in V_h$ , and  $\theta$  sufficiently small, there exists  $C > 0$ , independent of  $h$ , such that*

$$B_h(v, v) \geq C \|v\|^2. \quad (3)$$

**Proof** We first consider the term  $(\Delta v, \theta h\nu(v_t + \beta \cdot \nabla v))$ . On a single element  $\tau \in \mathcal{T}_h$ , we can use a local inverse assumption to see

$$|(-\Delta v, v_t + \beta \cdot \nabla v)_\tau| \leq C(h\nu)^{-1} \|\nabla v\|_\tau^2$$

Using this estimate, we have

$$\int_0^T \alpha(\nabla v, \nabla v) - (\Delta v, \theta h\nu(v_t + \beta \cdot \nabla v)) dt \geq \alpha(1 - C\theta) \int_0^T \|\nabla v\|^2 dt.$$

The term

$$\int_0^T (v_t + \beta \cdot \nabla v, v) dt = \frac{\|v(T)\|^2}{2}$$

and

$$(v_t + \beta \cdot \nabla v, \theta h\nu(v_t + \beta \cdot \nabla v)) = \theta h\nu \|v_t + \beta \cdot \nabla v\|^2.$$

Finally

$$\begin{aligned} (\gamma v, v + \theta h \nu(v_t + \beta \cdot \nabla v)) &\geq \gamma \|v\|^2 - \gamma \theta h \nu \|v\| \|v_t + \beta \cdot \nabla v\| \\ &\geq \gamma \|v\|^2 \left(1 - \frac{\gamma \theta h \nu}{2}\right) - \frac{\theta h \nu}{2} \|v_t + \beta \cdot \nabla v\|^2 \end{aligned}$$

Combining all these estimates, and taking  $\theta$  sufficiently small proves (3).  $\square$

The orthogonality-like relation for the error  $e = u - u_h$  in our approximation is given by

$$B_h(e, v) = 0 \quad (4)$$

for all  $v \in V_h$ .

For  $\chi \in V_h$ , let

$$\begin{aligned} \phi &= u_h - \chi \\ \eta &= u - \chi. \end{aligned}$$

Our error relation can be expressed in terms of  $\phi$  and  $\eta$  as

$$B_h(\phi, v) = B_h(\eta, v)$$

for all  $v \in V_h$ . We take  $v = \phi \in V_h$  and use Lemma 1. Then we have

$$\|\phi\|^2 \leq C B_h(\phi, \phi) \leq C B_h(\eta, \phi) \quad (5)$$

Let  $\delta$  be a sufficiently small parameter to be characterized below. We now estimate all the terms on the right hand side of (5). First,

$$\alpha(\nabla \eta, \nabla \phi) \leq C \alpha \|\nabla \eta\|^2 + \delta \alpha \|\nabla \phi\|^2$$

and

$$\alpha(\Delta \eta, \theta h \nu(\phi_t + \beta \cdot \nabla \phi)) \leq C \alpha h^2 \|\Delta \eta\|^2 + \delta \alpha \|\nabla \phi\|^2.$$

Next

$$\begin{aligned} (\eta_t + \beta \cdot \nabla \eta, \theta h \nu(\phi_t + \beta \cdot \nabla \phi)) &\leq C \theta h \nu \|\eta_t + \beta \cdot \nabla \eta\|^2 + \delta \theta h \nu \|\phi_t + \beta \cdot \nabla \phi\|^2 \\ (\gamma \eta, \phi + \theta h \nu(\phi_t + \beta \cdot \nabla \phi)) &\leq C \gamma \|\eta\|^2 + \delta (\gamma \|\phi\|^2 + \theta h \nu \|\phi_t + \beta \cdot \nabla \phi\|^2) \end{aligned}$$

The fifth term is a bit more involved.

$$\begin{aligned} \int_0^T (\eta_t + \beta \cdot \nabla \eta, \phi) dt &= (\eta(T), \phi(T)) - \int_0^T (\eta, \phi_t + \beta \cdot \nabla \phi) dt \\ &\leq C \left( \|\eta(T)\|^2 + (h \nu)^{-1} \int_0^T \|\eta\|^2 dt \right) \\ &\quad + \delta \left( \|\phi(T)\|^2 + h \nu \int_0^T \|\phi_t + \beta \cdot \nabla \phi\|^2 dt \right) \end{aligned}$$

Combining these five estimates, and making  $\delta$  sufficiently small, we have

$$\|\phi\|^2 \leq C \left( \|\eta\|^2 + \int_0^T (h\nu)^{-1} \|\eta\|^2 + \alpha h^2 \|\Delta\eta\|^2 dt \right) \quad (6)$$

Using (6) and the triangle inequality, we obtain

**Theorem 2.** *The error  $e = u - u_h$  satisfies*

$$\|u - u_h\|^2 \leq C \inf_{\chi \in V_h} \left( \|u - \chi\|^2 + \int_0^T (h\nu)^{-1} \|u - \chi\|^2 + \alpha h^2 \|\Delta(u - \chi)\|^2 dt \right) \quad (7)$$

Suppose  $\alpha = O(1)$  and  $V_h$  contains piecewise polynomials of degree  $r$ . Then if  $u \in \mathcal{H}^{r+1}(\Omega)$ , (7) yields an  $O(h^{r-1/2})$  rate of convergence of the space-time gradient in the streamline direction  $\mathbf{b} = (\beta^t, 1)^t$ , and an optimal  $O(h^r)$  convergence rate for the cross-wind direction(s). If  $\gamma = O(1)$  we have  $O(h^r)$  convergence for the  $\mathcal{L}_2$  norm. While not optimal in every norm considered, overall this is in alignment with well-known behavior for the classical streamline diffusion method. If  $\alpha = 0$  (or  $\alpha \ll O(h)$ ) we lose convergence in the cross-wind directions, but have improved  $O(h^r)$  convergence in the streamline direction, and if  $\gamma = O(1)$  we have improved  $O(h^{r+1/2})$  convergence for the  $\mathcal{L}_2$  norm. These again correspond with classical results for the streamline diffusion method.

## A Practical Remark

Suppose that the space domain  $\Omega_s$  has a generic length scale  $L$ . Since the time units for  $\Omega_t = [0, T]$  could be completely unrelated to the space units, the space-time domain  $\Omega = \Omega_s \times \Omega_t$  could be quite anisotropic. It could be very long if  $T \gg L$  or very short if  $T \ll L$ . Filling such potentially thin domains with a small number of shape regular elements could be problematic from the practical point of view. Therefore it could be useful to rescale the time variable such that it is has a similar scale to the space variables. For example, one could change variables as in

$$\tilde{t} = \frac{Lt}{T} \equiv \kappa t$$

for  $0 \leq \tilde{t} \leq L$ . The modified space time-domain  $\Omega_s \times [0, L]$  is more isotropic, and likely could be tessellated with far fewer shape regular elements. In terms of the partial differential equation,

$$\frac{\partial u}{\partial t} = \kappa \frac{\partial u}{\partial \tilde{t}}$$



making the convection in the time direction larger or smaller depending on the value of  $\kappa$ . In terms of our analysis, we could replace

$$\begin{aligned}\alpha &\rightarrow \frac{\alpha}{\kappa} \equiv \tilde{\alpha} \\ \beta &\rightarrow \frac{\beta}{\kappa} \equiv \tilde{\beta} \\ \gamma &\rightarrow \frac{\gamma}{\kappa} \equiv \tilde{\gamma}\end{aligned}$$

and directly apply the analysis of the previous section to this modified constant coefficient equation.

#### 4. High Order Scharfetter Gummel discretization

In this section we derive a high order Scharfetter-Gummel scheme on simplicial finite element grids in dimension  $d \geq 1$ . The original Scharfetter-Gummel difference scheme [10] is a method used in simulating 1-dimensional semiconductor equations. After its discovery, it has been generalized and used for the numerical solution of convection-diffusion equations of the form:

$$-\operatorname{div} J(u) = f, \quad x \in \Omega \subset \mathbb{R}^d \quad (8)$$

$$J(u) = (D(x)\nabla_x u - \mathbf{b}u), \quad (9)$$

$$u(x) = 0, \quad x \in \Gamma_D, \quad J(u) \cdot \mathbf{n} = 0, \quad x \in \Gamma_N \quad (10)$$

$$Du \cdot \mathbf{n} = 0, \quad x \in \Gamma_R. \quad (11)$$

Here,  $J(u)$  is the flux variable which plays an important role in approximating the weak form of the equation. We note that the natural boundary condition is the one given on  $\Gamma_N$  and the boundary condition (11) is a of Robin type for this problem. The weak form of the equation above is: Find  $u \in V$  such that

$$a(u, v) + m_R(u, v) = f(v), \quad (12)$$

$$a(u, v) = \int_{\Omega} J(u) \cdot \nabla v, \quad f(v) = \int_{\Omega} f v \quad (13)$$

$$m_R(u, v) = \int_{\Gamma_R} (\mathbf{b} \cdot \mathbf{n}) uv \quad (14)$$

The variational form is obtained after integration by parts and using the fact that on  $\Gamma_R$ ,  $Du \cdot \mathbf{n} = J(u) \cdot \mathbf{n} - \mathbf{b} \cdot \mathbf{n}u$ .

The Scharfetter-Gummel scheme was extended to more than 1 spatial dimension as the Edge Average Finite Element (EAFE) Scheme. A priori error estimates in any dimension were shown in [7]. This work only considered scalar valued diffusion coefficients (although in any spatial dimension); a discretization for matrix valued diffusion coefficients was proposed and analyzed in [9].

Here, we provide a novel approach which gives a Scharfetter-Gummel discretization for finite element spaces of order  $r \geq 1$ . Our approach follows the

ideas in [7] and [9]. The extension to  $r \geq 1$ , however is not at all straightforward and requires results from the recently developed Finite Element Exterior Calculus. The rationale of constructing the high order Scharfetter-Gummel scheme is:

- (i) approximate the flux  $J(u)$  via the Nédélec elements (discrete differential 1-forms with polynomial coefficients);
- (ii) eliminate the flux variable and write the resulting discrete problem in terms of the scalar valued finite element approximation of the solution of (1)  $u$  (a 0-form).

To set up the finite element approximation, we let us itemize some of the ingredients and the main assumptions needed for the discretization.

- We assume that  $\Omega$  is covered by a conforming, simplicial, shape-regular mesh  $\mathcal{T}_h$ . We have  $\Omega = \cup\{T \mid T \in \mathcal{T}_h\}$ .
- The space  $V_h$  is the space of conforming Lagrange finite elements of degree  $r$ . The crucial case is  $r = 1$ . For the derivation, we also need the Nédélec polynomial space on a fixed element (cf.eg. [11, 12, 13, 14]). The details are described below in §4.1.
- We assume that the flux  $J$  and  $u$  are smooth enough so that all the norms of functions below make sense. In particular  $J \in W^{1,p}(T)$ , for all  $T \in \mathcal{T}_h$  and for some  $p > d$ . The solution  $u$  is at least continuous, so that its Lagrange interpolant is well defined.
- We assume that the coefficients  $D$ ,  $\mathbf{b}$  are piece-wise constants with discontinuities aligned with  $\mathcal{T}_h$ .

Next, we show that (i) and (ii) in the rationale given earlier are computationally feasible steps.

#### 4.1. Notation and Nédélec spaces

Consider the Nédélec space  $P^\mathcal{N}$ , which restricted to any element  $T$  is the following polynomial space

$$\mathcal{P}^\mathcal{N} = (P_{r-1})^d \oplus \mathbf{S}_r, \quad P_r \supset P^\mathcal{N} \supset P_{r-1}, \quad (15)$$

where  $\mathbf{S}_r$  is the subspace of vector valued homogeneous polynomials of degree  $r$ , whose elements satisfy  $\mathbf{s} \cdot \mathbf{x} = 0$ . The inclusion relations on the right clearly hold by definition on the element  $T$ . From now on we fix this element. We refer to [11], [12], [15], [16], [17], for the classical and the modern description of these spaces and studies of their properties. In what follows we use some of the tools from [15] and [16]. In our notation, the lowest order of such polynomials corresponds to  $r = 1$ .

Further, let  $M = \dim P^\mathcal{N}$  be the dimension of the Nédélec polynomial space on  $T$ . The elements of the basis in the dual space of  $P_r^\mathcal{N}$  are known as *degrees*

of freedom and we denote them by  $\{\eta_j\}_{j=1}^M$ . Next, the basis in  $P^\mathcal{N}$ , dual to the degrees of freedom we denote by  $\{\varphi_j\}_{j=1}^M$ . For general simplex in  $\mathbb{R}^d$ , the explicit form of the degrees of freedom and their dual basis is found in [15]. For our purposes it is sufficient to note that the functionals  $\eta_j$  can be thought as integrals of traces of functions over sub-simplicies. For the lowest order case, we have

$$\langle \eta_e, \mathbf{v} \rangle = \int_e \mathbf{v} \cdot \boldsymbol{\tau}_e, \quad \varphi_e = \lambda_i \nabla \lambda_j - \lambda_j \nabla \lambda_i.$$

for every edge  $e = (i, j)$  of  $T$  (there are  $\frac{d(d+1)}{2}$  edges). Here,  $\boldsymbol{\tau}_e$  is the tangent for edge  $e$ , and  $\{\lambda_i\}$  are the usual baricentric coordinates for element  $T$  (cf., e.g., [18]). We also note the inclusions:

$$(P_{r-1})^d \subset P^\mathcal{N} \subset (P_r)^d$$

Using this notation, we have that any function  $\mathbf{v} \in P^\mathcal{N}$  can be written as

$$\mathbf{v} = \sum_{j=1}^M \langle \eta_j, \mathbf{v} \rangle \varphi_j(\mathbf{x}). \quad (16)$$

We stress that this representation is unique.

Such a representation provides a canonical interpolation operator, which for sufficiently smooth vector valued  $V$  is defined as

$$\Pi^\mathcal{N} \mathbf{v} = \sum_{j=1}^M \langle \eta_j, \mathbf{v} \rangle \varphi_j(\mathbf{x}). \quad (17)$$

The smoothness of  $\mathbf{v}$  must be such that the linear forms  $\langle \eta_j, \cdot \rangle$  are bounded.

Consider now the Lagrange finite elements of order  $r$ ,  $V_h$ . Similar definitions as above also exist for space  $V_h$ . That is, we have  $\mu_j$  the dual functionals, which in the present Lagrangian case, are simply pointwise evaluations. The Lagrangian basis is  $\{\xi_j\}$ . Explicit definitions are found in [18, Theorem 2.2.1].

Likewise, we have a canonical interpolation operator, well defined for any continuous  $v$ . The image of  $v \in C^0(\overline{\Omega})$  under this interpolation is denoted by  $v_I$  and we have

$$v_I = \sum_{j=1}^{N_h} \langle \mu_j, v \rangle \xi_j(\mathbf{x}). \quad (18)$$

There is no need to distinguish the global interpolation operator (on  $\Omega$ ) and the local one (on  $T \in \mathcal{T}_h$ ) for our considerations and we use the same notation for both. Let us note, however, that when working on fixed  $T \in \mathcal{T}_h$  we will use  $N = \dim P_r = \binom{r+d}{d}$ , instead of  $N_h = \dim V_h$ .

As is well known [15], we have commutative diagrams linking the Nédélec elements and the Lagrange elements of matching orders (order  $r$  here), and on every element  $T$  we have

$$\Pi^\mathcal{N} \nabla v = \nabla v_I.$$

This relation is in fact a relation between degrees of freedom, namely

$$\langle \eta_j, \nabla v \rangle = \langle \eta_j, \nabla v_I \rangle. \quad (19)$$

This is obvious by using the definition of  $\Pi^\mathcal{N}$ , the fact that  $\nabla v_I \in P^\mathcal{N}$ , and the uniqueness of the representation in (16).

#### 4.2. Derivation of a high order Scharfetter-Gummel scheme

Let us fix  $T \in \mathcal{T}_h$  and we start with the definition of  $J$  and use that  $D$  and  $\mathbf{b}$  are constant on  $T$ .

$$J(u) = D\nabla u - \mathbf{b}u = \exp(\mathbf{q} \cdot \mathbf{x}) D \nabla (\exp(-\mathbf{q} \cdot \mathbf{x}) u), \quad \mathbf{q} = D^{-1} \mathbf{b}.$$

Hence, we have

$$\exp(-\mathbf{q} \cdot \mathbf{x}) D^{-1} J(u) = \nabla (\exp(-\mathbf{q} \cdot \mathbf{x}) u). \quad (20)$$

If we apply now  $\langle \eta_j, \cdot \rangle$ ,  $j = 1 : M$  on both sides, and then use (19) we get

$$\begin{aligned} \langle \eta_j, e^{(-\mathbf{q} \cdot \mathbf{x})} D^{-1} J(u) \rangle &= \langle \eta_j, \nabla (e^{(-\mathbf{q} \cdot \mathbf{x})} u) \rangle \\ &= \langle \eta_j, \nabla (e^{(-\mathbf{q} \cdot \mathbf{x})} u)_I \rangle, \\ &= \langle \eta_j, \nabla (e^{(-\mathbf{q} \cdot \mathbf{x})} u_I)_I \rangle, \end{aligned}$$

The latter identity on the right hand side above, uses the fact that the " $I''$ -interpolant is based on the functionals  $\mu_j$  that are based on nodal evaluation. As expected, the right hand side is a gradient of a function in  $V_h$  and in summary we have

$$\langle \eta_j, e^{(-\mathbf{q} \cdot \mathbf{x})} D^{-1} J(u) \rangle = \langle \eta_j, \nabla (e^{(-\mathbf{q} \cdot \mathbf{x})} u_I)_I \rangle, \quad j = 1, \dots, M. \quad (21)$$

Introducing now  $\mathbf{G}(J(u)) \in \mathbb{R}^M$  and  $\mathbf{d}(u) \in \mathbb{R}^M$  by

$$[\mathbf{G}(J(u))]_j = \langle \eta_j, e^{(-\mathbf{q} \cdot \mathbf{x})} D^{-1} J(u) \rangle, \quad j = 1, \dots, m \quad (22)$$

$$[\mathbf{d}(u)]_j = \langle \eta_j, \nabla (e^{(-\mathbf{q} \cdot \mathbf{x})} u_I)_I \rangle, \quad j = 1, \dots, M. \quad (23)$$

and we can write (21) as

$$\mathbf{G}(J(u)) = \mathbf{d}(u), \quad (24)$$

Note that both  $\mathbf{G}$  and  $\mathbf{d}$  are linear operators, mapping vector fields and functions to  $\mathbb{R}^M$ . We remark that the relation (24) is used later in the definition of the approximate bilinear form, and, in the proof of the error estimates, and, we further stress on the fact that  $\mathbf{d}(u) = \mathbf{d}(u_I)$ , by definition.

The main idea of the Scharfetter-Gummel and EAFE schemes is to approximate  $J(u)$ ,

$$J(u) \approx J_T(u) \in P^\mathcal{N},$$

or equivalently, we seek

$$J_T(u) = \sum c_j \varphi_j,$$

for some coefficient vector  $\mathbf{c} = (c_j)$ . The coefficient  $\mathbf{c}$  is chosen so that the relation (21) still holds for the approximation. An important question is whether this is possible. If  $r = 1$ , and we use the lowest order Nédélec elements, this is definitely the case as shown in the earlier works [7, 9].

A construction with higher order polynomial spaces is a bit more intricate. In general, we would like to find  $J_T(u) \in P^\mathcal{N}$ . A key observation is that in order to derive our scheme, we use the weak form of the equation (8) and we will aim to approximate the weak form as follows:

$$\int_T J(u_h) \cdot \nabla v_h \approx \int_T J_T(u_h) \cdot \nabla v_h,$$

for functions  $v_h \in V_h$  and  $J_T(u_h) \approx J(u_h)$ . As on  $T$ ,  $\nabla v_h \in (P_{r-1})^d$ , it is sufficient to look for approximations  $J_T(u) \in (P_{r-1})^d \subset P^\mathcal{N}$ .

We now explore this observation and look at how it affects the identity (24). Let  $P$  be the matrix representation of the embedding  $(P_{r-1})^d \subset P^\mathcal{N}$ . To define this matrix, let  $\{\psi_m\}_{m=1}^{M_0}$  be a basis in  $(P_{r-1})^d$ , with  $M_0 = \dim(P_{r-1})^d$  and let  $\{\varphi_j\}_{j=1}^M$  be the basis in  $P^\mathcal{N}$ , dual to the degrees of freedom  $\{\eta_m\}_{m=1}^M$ . Then the entries of  $P$  are the coefficients in the expansion of  $\psi_m$  in terms of  $\{\varphi_j\}_{j=1}^M$ , and we have,

$$\psi_k = \sum_{j=1}^M p_{jk} \varphi_j, \quad \text{with} \quad p_{mk} = \langle \eta_m, \psi_k \rangle. \quad (25)$$

Note that  $\nabla(\exp(-\mathbf{q} \cdot \mathbf{x})u_I)_I$  is an element  $(P_{r-1})^d$ , and, as such, it can be written as a linear combination via  $\{\psi_k\}_{k=1}^{M_0}$ . Recalling the definition of  $\mathbf{d}(u)$  in (23) then leads to the following useful relations:

$$\begin{aligned} \nabla(\exp(-\mathbf{q} \cdot \mathbf{x})u_I)_I &= \sum_{k=1}^{M_0} \tilde{d}_k \psi_k, \quad [\mathbf{d}(u)]_j = \sum_{k=1}^{M_0} \langle \eta_j, \psi_k \rangle \tilde{d}_k, \\ \mathbf{d}(u) &= P \tilde{\mathbf{d}}. \end{aligned}$$

As a consequence, to define the approximation  $J_T$ , we need to find a solution of the following problem

$$P^* Z P \tilde{\mathbf{c}} = P^* P \tilde{\mathbf{d}}, \quad (26)$$

where we have set  $\mathbf{c} = P \tilde{\mathbf{c}}$ , and, as we have shown,  $\mathbf{d} = \tilde{\mathbf{d}}$ . Above the matrix  $Z \in \mathbb{R}^{M \times M}$  has entries

$$Z_{jk} = \langle \eta_j, e^{-\mathbf{b} \cdot D^{-1} \mathbf{x}} D^{-1} \varphi_k \rangle.$$

The following remark is in order. In general, we may have tried to solve the following system of equations for the coefficients  $\mathbf{c}$ :

$$Z \mathbf{c} = \mathbf{d}. \quad (27)$$

Clearly, if this is a well posed problem, then we can find the approximation  $J_T(u)$ . However, as it described above, we only need the solution in subspace,

i.e., to solve problem (26). To show that the subspace problem (26) is solvable it is sufficient to show that  $ZP$  is injective, and since  $P$  is injective, it is sufficient to show that  $Z$  is injective on the range of  $P$ .

We let

$$Z^\dagger = P(P^* Z P)^{-1} P^*. \quad (28)$$

In the following, we will simply denote  $\mathbf{c} = P\tilde{\mathbf{c}}$  by  $\mathbf{c} = Z^\dagger \mathbf{d}$ , or, by (24), by  $Z^\dagger G(J(u))$ .

The following lemma follows by the construction of the approximation  $J_T(u)$ .

**Lemma 3.** *If  $J(u)$  is polynomial of degree  $r - 1$ , then its approximation  $J_T(u)$  defined by  $Z^\dagger G(J(u))$  coincides with  $J(u)$ .*

#### 4.2.1. A unisolvence result

Note that, when  $\mathbf{b} = 0$ , the solvability of such system follows from the fact that the Nédélec degrees of freedom form a unisolvent set of functionals on  $P^N \supset P_{r-1}\Lambda^1(T)$ . Multiplying by the exponent changes the game, and, we need to prove some of the basic results on unisolvence of Nédélec degrees of freedom for quasi-polynomials which we state in the following lemma.

**Lemma 4.** *The matrix  $ZP$  is injective, or, equivalently, if  $\mathbf{u} \in (P_{r-1})^d$  and  $\langle \eta_j, e^{(-\mathbf{b} \cdot D^{-1} \mathbf{x})} \mathbf{p} \rangle = 0$ , for all  $j = 1 : M$ , then  $\mathbf{u} = 0$ .*

**Proof** This proof follows exactly the lines of the proofs of [16, Lemma 4.5, Lemma 4.6]. The only modifications needed are that we multiply by an exponential function, which is positive everywhere. The rest of the arguments carry over without any change.  $\square$

#### 4.2.2. Derivation of the discrete problem

Since now the approximation  $J_T(u_h)$ , for  $u_h \in V_h$  is well defined, due Lemma 4, we have a natural approximating bilinear form. For  $u_h \in V_h$  and  $v_h \in V_h$  we set

$$\begin{aligned} a_h(u_h, v_h) &= \sum_T \int_T J_T(u_h) \cdot \nabla v_h \\ &= \sum_T \sum_{j=1}^M \int_T [Z_T^\dagger d_T(u_h)]_j \int_T \varphi_j \cdot \nabla v_h. \end{aligned} \quad (29)$$

The coefficients in  $J_T(u_h)$  are determined by  $Z_T^\dagger \mathbf{d}_T(u_h)$ , for all  $T \in \mathcal{T}_h$ , which in turn indeed makes the right side of (29) to depend only on the degrees of freedom of  $u_h$ .

We then define the following discrete problem: Find  $u_h \in V_h$  such that

$$a_h(u_h, v) = f(v), \quad \text{for all } v \in V_h. \quad (30)$$

We note another useful relation which follows from the derivation above and is used in the error estimates below. It is an analogue of [7, Equation (3.16)],

and [9, Equation (3.8)] and it plays a crucial role in the a priori error estimates. In particular it is useful to estimate the deviation of the derived discrete scheme from the standard Galerkin one (with bilinear form  $a(u, v) = \int_{\Omega} (D\nabla u - \mathbf{b}u) \cdot \nabla v$ ).

**Lemma 5.** *For any continuous  $u$ , and sufficiently smooth  $J$ , such that  $\Pi^N J(u)$  is well defined we have:*

$$a_h(u_I, v_h) = \sum_T \sum_{j=1}^M [Z_T^\dagger G_T(J(u))]_j \varphi_j \cdot \nabla v_h. \quad (31)$$

**Proof** Then the relation follows by recalling that  $d_T(u_I) = d_T(u)$ , and substituting (24) in (29).  $\square$

Next, we show that, under certain conditions, this is a well posed problem, and we also prove an *a priori* error estimate.

#### 4.3. Stability and error analysis

Error estimates and other properties of such discretization schemes are found in [7], [9]. Here we give an estimate for higher order Scharfetter-Gummel discretization and assume for simplicity, and without loss of any generality that we have Dirichlet boundary conditions. We have the following theorem:

**Theorem 6.** *Assume that  $a(\cdot, \cdot)$  is invertible on  $V_h$ . Then, for sufficiently small  $h$ , the discrete variational problem (30) is well posed and the following error estimate holds:*

$$|u_I - u_h|_{1,\Omega} \leq ch^r |J(u)|_{r,p,\Omega}. \quad (32)$$

**Proof** From the definition of  $a_h(\cdot, \cdot)$ , for all  $v \in V_h$  we have

$$\begin{aligned} |a(u, v) - a_h(u_I, v)| &= \left| \sum_T a_T(u, v) - a_{h,T}(u_I, v) \right| \\ &\leq \sum_T \left| \int_T J(u) \cdot \nabla v - \sum_{j=1}^M [Z_T^\dagger G_T(J(u))]_j \int_T \varphi_j \cdot \nabla v \right|. \end{aligned}$$

Note that from Lemmas 5 and 3, the right side vanishes for all  $J(u)$  that are polynomials of degree  $(r-1)$ , and, standard scaling argument shows the estimate

$$|a(u, v) - a_h(u_I, v)| \leq ch^r |J(u)|_{r,p,\Omega} |v|_{1,q,\Omega}, \quad p^{-1} + q^{-1} = 1. \quad (33)$$

The solvability of the discrete problem then follows from the fact that by assumption  $a(\cdot, \cdot)$  provides a solvable problem, and hence it satisfies an inf-sup condition on  $V_h$ . According to (33)  $a(\cdot, \cdot)$  and  $a_h(\cdot, \cdot)$  are close when  $h \rightarrow 0$ , and, hence,  $a(\cdot, \cdot)$  also satisfies an inf-sup condition for sufficiently small  $h$ . This in turn implies that the discrete problem (30) is well posed. The error estimate (32) then follows from the inequality (33). We refer to [7, Lemma 6.2 and Theorem 6.3] for all the details missing here.  $\square$

## 5. Application to parabolic problems

In this section, we recall the parabolic equation (1):

$$\begin{aligned} u_t - \operatorname{div}(K(x)\nabla_x u - \beta u) &= f, \\ u(x, 0) &= u_0(x), \quad \text{for } t = 0; \\ u(x, t) &= 0, \quad x \in \Gamma = \partial\Omega \times \{[0, t_{\max}]\}. \end{aligned} \quad (34)$$

This equation and the equation discretized by the streamline diffusion method match, if  $\operatorname{div}_x \beta = 0$ , which we assume to hold. In general, the divergence form comes from a material law and many mathematical models of physical phenomena (if not all) are in divergence form.

The space-time formulation, (written in terms of a flux  $J_0$ , and with  $y = (x, t)$ ) then is:

$$\begin{aligned} -\operatorname{div}_y J_0(u) &= f, \quad \tilde{J}_0(u) = D_0(x)\nabla_x u - \mathbf{b}u \\ u(y) &= 0, \quad x \in \Gamma = \partial\Omega \times \{(0, t_{\max}]\}, \\ u(y) &= u_0(x), \quad x \in \Gamma_0 = \overline{\Omega} \times \{t = 0\}. \end{aligned} \quad (35)$$

Here we have introduced the semidefinite, tensor valued function  $D_0(x)$ , and more generally, we denote,  $D_\varepsilon(x) : \Omega \mapsto \mathbb{R}^{(d+1) \times (d+1)}$ :

$$D_\varepsilon = \begin{pmatrix} K(x) & 0 \\ 0 & \varepsilon \end{pmatrix}, \quad J_\varepsilon = D_\varepsilon \nabla_y u - \mathbf{b}u. \quad (36)$$

The well known heat equation,  $u_t - \Delta u = f$ , corresponds to  $K(x) = I$  and  $\beta = 0$  and  $\varepsilon = 0$ .

The technique described in the previous section does not work in a straightforward fashion in the case of space-time formulation, because  $D_0$  is a singular matrix. In fact, there is no obvious construction that works in the case of singular  $D_0$ . We consider then a formulation using perturbation of the diffusion tensor  $D_\varepsilon$  and the flux  $J_\varepsilon$ . Thus, for the parabolic problem we set

$$J_\varepsilon(u) = D_\varepsilon \nabla_y u - \mathbf{b}u, \quad \mathbf{b} = (\beta_K, 1)^T.$$

### 5.1. Lowest order discretization for parabolic equations

In this section we discuss the Scharfetter-Gummel discretization when applied to space-time formulation of a parabolic equation, in the lowest order case. As a simple, but important example, we consider the simple case of heat equation, i.e.  $\beta_K = 0$ , which implies that  $\mathbf{b} = \mathbf{e}_{d+1}$ , and  $\mathbf{e}_{d+1} = \underbrace{(0, \dots, 0)}_d, 1)^T$ .

We next compute the action of the local stiffness matrix corresponding to a parabolic problem on a vector of degrees of freedom  $u \in V_h$ . We fix an element  $((d+1)$  dimensional simplex)  $T \in \mathcal{T}_h$  and we denote its barycentric coordinates by  $\{\lambda_i\}_{i=1}^{d+2}$  and the space-time coordinates of its vertices are  $\{\mathbf{y}_i\}_{i=1}^{d+2} = \{(\mathbf{x}_i, t_i)\}_{i=1}^{d+2}$ . The degrees of freedom of a linear polynomial  $u \in V_h$



restricted to  $T$  are  $\{u_i\}_{i=1}^{d+2} = \{u(\mathbf{y}_i)\}_{i=1}^{d+2}$  and we have  $u(\mathbf{y}) = \sum_{i=1}^{d+2} u_i \lambda_i(\mathbf{y})$ . For an edge  $E \in T$ ,  $E = (\mathbf{y}_i, \mathbf{y}_j)$ ,  $i = 1, \dots, (d+2)$ ,  $j = 1, \dots, (d+2)$ , we denote

$$\boldsymbol{\tau}_{ij} = \boldsymbol{\tau}_E = \frac{(\mathbf{y}_i - \mathbf{y}_j)}{|\mathbf{y}_i - \mathbf{y}_j|}, \quad |\mathbf{r}| = \sqrt{\sum_{l=1}^{d+1} r_l^2}, \quad \text{for all } \mathbf{r} \in \mathbb{R}^{d+1}.$$

We note that  $\boldsymbol{\tau}_{ij} = -\boldsymbol{\tau}_{ji}$ , but as we shall see, this is of no consequence for the final form of the local stiffness matrix. To avoid complications in the presentation coming from unnecessary subscripts we will write  $D$  (resp.  $J$ ) instead of  $D_\varepsilon$  and (resp.  $J_\varepsilon$ ).

For any  $u \in V_h$ , as  $D$  and  $J(u)$  are constants on  $T$ , we have the following obvious identities from the definition of  $J$ :

$$\begin{aligned} D^{-1} J \cdot \boldsymbol{\tau}_E &= e^{t/\varepsilon} \nabla_{\mathbf{y}} \left( e^{-t/\varepsilon} u \right), \\ \int_E e^{-t/\varepsilon} D^{-1} J \cdot \boldsymbol{\tau}_E dE &= \int_E \nabla_{\mathbf{y}} \left( e^{-t/\varepsilon} u \right) \cdot \boldsymbol{\tau}_E dE, \\ (D^{-1} J \cdot \boldsymbol{\tau}_E) \int_E e^{-t/\varepsilon} dE &= [e^{-t_i/\varepsilon} u(\mathbf{x}_i, t_i) - e^{-t_j/\varepsilon} u(\mathbf{x}_j, t_j)]. \end{aligned}$$

Computing the integral on the left side gives

$$\begin{aligned} \int_E e^{-\frac{t}{\varepsilon}} dE &= |E| \int_0^1 \exp \left( -\frac{t_j + s(t_i - t_j)}{\varepsilon} \right) ds \\ &= \frac{|E|\varepsilon}{t_j - t_i} \int_{-t_j/\varepsilon}^{-t_i/\varepsilon} e^\xi d\xi = |E|\varepsilon \frac{e^{-t_i/\varepsilon} - e^{-t_j/\varepsilon}}{t_j - t_i} \\ &= \frac{|E|e^{-t_i/\varepsilon}}{B(\frac{t_i - t_j}{\varepsilon})} = \frac{|E|e^{-t_j/\varepsilon}}{B(\frac{t_j - t_i}{\varepsilon})}, \end{aligned}$$

where  $B(s) = \frac{s}{e^s - 1}$  is the Bernoulli function ( $B(0) = 1$ ). Note that,  $B(s) = e^{-s} B(-s)$  and  $B(s) > 0$  for all  $s \in \mathbb{R}$ . We then conclude that on every edge  $E$  in  $T$  we have:

$$|E|(D^{-1} J \cdot \boldsymbol{\tau}_E) = B \left( \frac{t_i - t_j}{\varepsilon} \right) u(\mathbf{y}_i) - B \left( \frac{t_j - t_i}{\varepsilon} \right) u(\mathbf{y}_j). \quad (37)$$

In the derivation for general order of polynomials we needed the Nédélec basis and spaces. In the lowest order case, we can take a route that does not use these spaces explicitly. In the evaluation of the stiffness matrix entries, we need to compute integrals of the form

$$\int_T (J \cdot \nabla_{\mathbf{y}} \lambda_j) = |T|(J \cdot \nabla_{\mathbf{y}} \lambda_j).$$

We note that since  $D^{-1} J$  is a constant on  $T$ , we can write it as a gradient of a linear function, namely

$$J = D(D^{-1} J) = D \nabla_{\mathbf{y}} (D^{-1} J \cdot \mathbf{y}) = \sum_{i=1}^{d+2} (D^{-1} J \cdot \mathbf{y}_i) D \nabla_{\mathbf{y}} \lambda_i. \quad (38)$$

Since  $\sum_{i=1}^{d+2} \nabla_y \lambda_i \equiv 0$  on  $T$ , we have that

$$0 = (D^{-1} J \cdot \mathbf{y}_j) \left( D \sum_{i=1}^{d+2} \nabla_y \lambda_i \cdot \nabla_y \lambda_j \right) = \sum_{i=1}^{d+2} (D^{-1} J \cdot \mathbf{y}_j) (D \nabla_y \lambda_i \cdot \nabla_y \lambda_j)$$

Hence,

$$\begin{aligned} |T|(J \cdot \nabla_y \lambda_j) &= \sum_{i=1}^{d+2} (D^{-1} J \cdot \mathbf{y}_i) (D \nabla_y \lambda_i \cdot \nabla_y \lambda_j) \\ &= \sum_{i \neq j} (D^{-1} J \cdot (\mathbf{y}_i - \mathbf{y}_j)) (D \nabla_y \lambda_i \cdot \nabla_y \lambda_j) \\ &= \sum_{i \neq j} |E| (D^{-1} J \cdot \boldsymbol{\tau}_{ij}) (D \nabla_y \lambda_i \cdot \nabla_y \lambda_j). \end{aligned}$$

We have computed earlier the quantity  $|E|(D^{-1} J \cdot \boldsymbol{\tau}_{ij})$  for all  $E \subset \partial T$ . Therefore,

$$|T|(J \cdot \nabla_y \lambda_j) = \sum_{i=1; i \neq j}^{d+2} d_{ji}^T \left[ B \left( \frac{t_i - t_j}{\varepsilon} \right) u_i - B \left( \frac{t_j - t_i}{\varepsilon} \right) u_j \right]. \quad (39)$$

Here  $d_{ji}^T = \int_T D \nabla_y \lambda_i \cdot \nabla_y \lambda_j$  are the entries of the local stiffness matrix corresponding to the discretization of  $(-\operatorname{div} D \nabla)$  with linear elements on  $T$ . Therefore on  $T$  we get

$$[A_T]_{jj} = - \sum_{i=1; i \neq j}^{d+2} d_{ji}^T B \left( \frac{t_j - t_i}{\varepsilon} \right), \quad [A_T]_{ji} = d_{ji}^T B \left( \frac{t_i - t_j}{\varepsilon} \right). \quad (40)$$

The global stiffness matrix is assembled from  $A_T$ . It is invertible for sufficiently small mesh size, invertible whenever the assembly of  $d_{ij}^T$  gives an  $M$ -matrix.

For more detailed discussions about sufficient conditions which lead to a stiffness matrix which is an  $M$ -matrix, as well as relations to finite volume methods we refer to [8]. More importantly, the work [8] provides techniques for consistent modification of the local stiffness matrices, leading to solvable linear systems for wide range of meshes (not only Delaunay meshes, or meshes satisfying the condition given in [7, Lemma 2.1]).

## 6. Numerical tests

We consider the 2D heat equation with Dirichlet boundary conditions on the unit square  $(0, 1) \times (0, 1)$ . We test both schemes: StreamLineDiffusion (SLDI) and EAFE on a uniform triangulation of the unit square. The exact solution is

$$U(x, t) = e^{-t} \sin \pi x \sin \pi y.$$

The domain is the unit square in 2D, and the space-time problem is solved as fully coupled 3D convection diffusion problem.

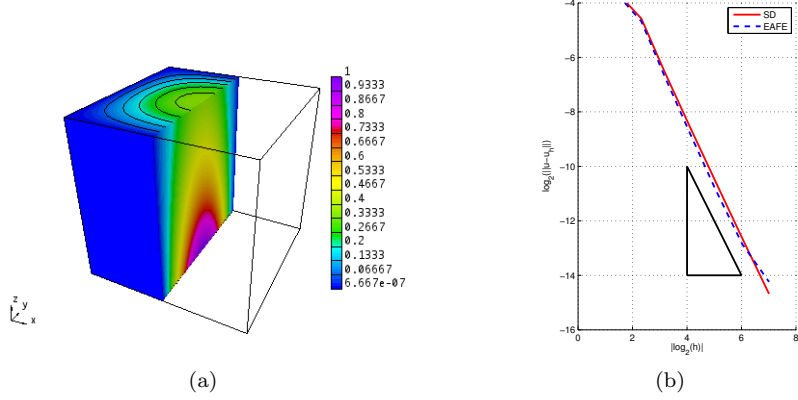


Figure 1: (a) Trace of the solution on the plane  $x = \frac{1}{2}$ . (b) Error reduction in  $L^2$ -norm. Quadratic convergence is clearly observed.

We have tested the lowest order streamline diffusion scheme which has the same number of degrees of freedom as the EAFE scheme. The convergence behavior of both discretizations is shown in Figure 1b.

We have tested the convergence on a family of successively refined triangulations. The coarsest one has a mesh size  $h_0 \approx \frac{1}{2}$  and the finest  $2^{-8}$  in 3D. The parameter  $\varepsilon$  in the diffusion tensor  $D_\varepsilon$  for the EAFE scheme was  $10^{-5}$  on all grids. The parameter  $\theta$  in the streamline diffusion method was set to  $10^{-2}$  on all grids. Such pool of tests corresponds to mesh with 27 vertices on the coarsest grid, and,  $\approx 2.1 \times 10^6$  vertices on the finest grid. In Figure 1a we have plotted the trace of the approximate solution on the plane  $x = 0$ . The approximate solution obtained via the EAFE scheme looks exactly the same, as is also the exact solution.

We next show a plot of a solution to an equation with convection depending on time. The equation is

$$u_t - \operatorname{div}(K(x)\nabla u - \mathbf{b}u) = 1, \quad x \in \Omega_s, \quad \mathbf{b} = \begin{pmatrix} 100 \sin(6\pi t) \\ 0 \end{pmatrix} \quad (41)$$

and the boundary and initial conditions are homogeneous, the domain is the unit square and the time interval is  $(0, 1)$ . The solution via the Scharfetter-Gummel (EAFE) scheme is shown in Figure 2. Note that with such convection term, the convection is 0 for  $t = k/6$  and  $k$  integer; it is, however, convection dominated for other values of  $t$ .

We have mentioned already the software used in performing the tests. In summary, we have used the C++ library and examples from the `mfem` package [19] (discretization); The solutions of the resulting linear systems are done using the Algebraic Multigraph Multilevel ILU algorithm by Bank and Smith [20, 21] found at <http://ccom.ucsd.edu/~reb/software.html>. the visualization was done using the `glvis` tool [22].

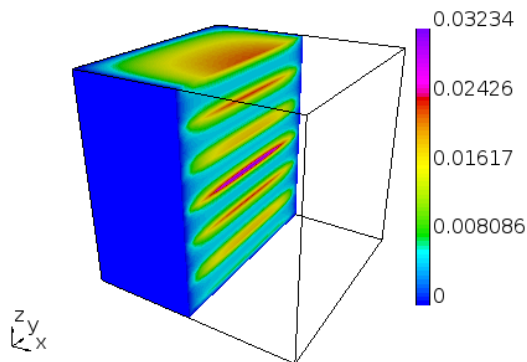


Figure 2: Trace of the numerical solution of equation (41) on the plane  $x = \frac{1}{2}$ . The effect of the time dependent convection is clearly seen in the plot.

- [1] I. V. Lashuk, P. S. Vassilevski, The Construction of coarse de Rham complexes with improved approximation properties, Computational Methods in Applied Mathematics 14(2) (2014) 257–303, URL <http://www.degruyter.com/view/j/cmam.2014.14.issue-2/cmam-2014-0004/cmam-2014-0004.xml>.
- [2] M. Griebel, D. Oeltz, P. Vassilevski, Space-time approximation with sparse grids, SIAM Journal on Scientific Computing 28 (2006) 701–727.
- [3] U. Langer, S. Moore, M. Neumüller, Space-time isogeometric analysis of parabolic evolution equations, RICAM Report 19, Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, also available at <http://arxiv.org/abs/1509.02008>, 2015.
- [4] A. N. Brooks, T. J. R. Hughes, Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations, Comput. Methods Appl. Mech. Engrg. 32 (1-3) (1982) 199–259, ISSN 0045-7825, URL [http://dx.doi.org/10.1016/0045-7825\(82\)90071-8](http://dx.doi.org/10.1016/0045-7825(82)90071-8), fENOMECH '81, Part I (Stuttgart, 1981).
- [5] C. Johnson, Numerical solution of partial differential equations by the finite element method, Cambridge University Press, Cambridge, ISBN 0-521-34514-6; 0-521-34758-0, 1987.
- [6] O. Axelsson, V. Eijkhout, B. Polman, P. Vassilevski, Incomplete block-matrix factorization iterative methods for convection-diffusion problems, BIT 29 (4) (1989) 867–889, ISSN 0006-3835, URL <http://dx.doi.org/10.1007/BF01932750>.

- [7] J. Xu, L. Zikatanov, A monotone finite element scheme for convection-diffusion equations, *Math. Comp.* 68 (228) (1999) 1429–1446.
- [8] R. E. Bank, W. Coughran, Jr., L. C. Cowsar, The Finite Volume Scharfetter-Gummel method for steady convection diffusion equations, *Computing and Visualization in Science* 1 (3) (1998) 123–136, ISSN 1432-9360, URL <http://dx.doi.org/10.1007/s007910050012>.
- [9] R. D. Lazarov, L. T. Zikatanov, An exponential fitting scheme for general convection-diffusion equations on tetrahedral meshes, *Computational and Applied Mathematics (Obchysljuval'na ta prykladna matematyka, Kiev)* 92 (1) (2005) 60–69, URL <http://isc.tamu.edu/publications-reports/#2004>, also available as Tech. Report 04-15, Institute for Scientific Computations, Texas A&M University, 2004.
- [10] D. Scharfetter, H. Gummel, Large-signal analysis of a silicon read diode oscillator, *IEEE Trans. Electron Devices* ED-16 (205) (1969) 959–962.
- [11] J.-C. Nédélec, Mixed finite elements in  $\mathbf{R}^3$ , *Numer. Math.* 35 (3) (1980) 315–341, ISSN 0029-599X, URL <http://dx.doi.org/10.1007/BF01396415>.
- [12] J.-C. Nédélec, A new family of mixed finite elements in  $\mathbf{R}^3$ , *Numer. Math.* 50 (1) (1986) 57–81, ISSN 0029-599X, URL <http://dx.doi.org/10.1007/BF01389668>.
- [13] A. Bossavit, A rationale for edge elements in 3D field computations,, *IEEE Trans. Mag.* 24 (1988) 74–79.
- [14] P. Monk, *Finite element methods for Maxwell's equations*, Numerical Mathematics and Scientific Computation, Oxford University Press, New York, ISBN 0-19-850888-3, URL <http://dx.doi.org/10.1093/acprof:oso/9780198508885.001.0001>, 2003.
- [15] R. Hiptmair, Canonical construction of finite elements, *Math. Comp.* 68 (228) (1999) 1325–1346, ISSN 0025-5718, URL <http://dx.doi.org/10.1090/S0025-5718-99-01166-7>.
- [16] D. N. Arnold, R. S. Falk, R. Winther, Finite element exterior calculus, homological techniques, and applications, *Acta Numer.* 15 (2006) 1–155, ISSN 0962-4929, URL <http://dx.doi.org/10.1017/S0962492906210018>.
- [17] D. N. Arnold, R. S. Falk, R. Winther, Finite element exterior calculus: from Hodge theory to numerical stability, *Bull. Amer. Math. Soc. (N.S.)* 47 (2) (2010) 281–354, ISSN 0273-0979, URL <http://dx.doi.org/10.1090/S0273-0979-10-01278-4>.
- [18] P. G. Ciarlet, *The finite element method for elliptic problems*, vol. 40 of *Classics in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, ISBN 0-89871-514-8, URL <http://>

[//dx.doi.org/10.1137/1.9780898719208](http://dx.doi.org/10.1137/1.9780898719208), reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)], 2002.

- [19] Software, MFEM: Modular finite element methods, <http://mfem.org>, 2015.
- [20] R. E. Bank, R. K. Smith, An algebraic multilevel multigraph algorithm, SIAM J. Sci. Comput. 23 (5) (2002) 1572–1592 (electronic), ISSN 1064-8275, URL <http://dx.doi.org/10.1137/S1064827500381045>.
- [21] R. E. Bank, Multigraph 1.0: An algebraic multilevel multigraph algorithm, <http://ccom.ucsd.edu/~reb/software.html>, 2015.
- [22] Software, GLVis: Accurate finite element visualization, <http://glvis.org>, 2015.